

CLAIMS

What is claimed is:

1. A method for debugging and tuning synthesized audio, comprising the steps of:
displaying a waveform corresponding to synthesized audio generated from concatenated phonetic units;
displaying parameters corresponding to at least one of the phonetic units;
displaying original recording containing selected phonetic unit;
receiving an editing input from the user; and
adjusting the parameters in accordance with the editing input.
2. The method of claim 1, wherein said displaying parameters step further comprises automatically displaying the parameters responsive to a user selection of at least a portion of the waveform, the displayed parameters correlating to the selected portion of the waveform.
3. The method of claim 1, wherein said displaying parameters step further comprises identifying a portion of the waveform responsive to a user selection of at least one of the parameters, the identified portion of the waveform correlating to the selected parameters.
4. The method of claim 1, wherein the edited parameters are contained in a text-to-speech engine configuration file.
5. The method of claim 4, the edited parameters comprising at least one property selected from the group consisting of speed, base pitch, volume, and search cost function weights.
6. The method of claim 1, wherein the edited parameters are contained in a segment dataset.

7. The method of claim 5, wherein the parameters comprise at least one parameter selected from the group consisting of a phonetic unit label, a phonetic unit boundary, a pitch mark and a phonetic alignment.
8. The method of claim 5, wherein said editing step comprises at least one action selected from the group consisting of deleting a pitch mark, inserting a pitch mark, repositioning a pitch mark and adjusting a phonetic alignment.
9. The method of claim 5, wherein said automatically displaying parameters step further comprises the step of displaying a recording's waveform associated containing the phonetic unit.
10. The method of claim 9, wherein edits to the waveform adjust parameters in the segment dataset.
11. The method of claim 1, wherein the synthesized audio is generated from a text input.
12. The method of claim 10, wherein the text input is received from the user.
13. A machine-readable storage having stored thereon a computer program having a plurality of code sections, the code sections executable by a machine for causing the machine to perform the steps of:
 - displaying a waveform corresponding to synthesized audio generated from concatenated phonetic units;
 - displaying parameters corresponding to at least one of the phonetic units;
 - displaying original recording containing selected phonetic unit;
 - receiving an editing input from the user; and
 - adjusting the parameters in accordance with the editing input.

14. The machine-readable storage of claim 13, wherein said displaying parameters step further comprises automatically displaying the parameters responsive to a user selection of at least a portion of the waveform, the displayed parameters correlating to the selected portion of the waveform.

15. The machine-readable storage of claim 13, wherein said displaying parameters step further comprises identifying a portion of the waveform responsive to a user selection of at least one of the parameters, the identified portion of the waveform correlating to the selected parameters.

16. The machine-readable storage of claim 13, wherein the edited parameters are contained in a text-to-speech engine configuration file.

17. The machine-readable storage of claim 16, the edited parameters comprising at least one property selected from the group consisting of speed, base pitch, volume, and search cost function weights.

18. The machine-readable storage of claim 13, wherein the edited parameters are contained in a segment dataset.

19. The machine-readable storage of claim 18, wherein the parameters comprise at least one parameter selected from the group consisting of a phonetic unit label, a phonetic unit boundary, a pitch mark and a phonetic alignment.

20. The machine-readable storage of claim 18, wherein said editing step comprises at least one action selected from the group consisting of deleting a pitch mark, inserting a pitch mark, repositioning a pitch mark and adjusting a phonetic alignment.

21. The machine-readable storage of claim 18, wherein said automatically displaying parameters step further comprises the step of displaying a recording's waveform associated containing the phonetic unit.
22. The machine-readable storage of claim 21, wherein edits to the waveform adjust parameters in the segment dataset.
23. The machine-readable storage of claim 13, wherein the synthesized audio is generated from a text input.
24. The machine-readable storage of claim 23, wherein the text input is received from the user.
25. A method for debugging and tuning synthesized audio, comprising the steps of:
 - means for displaying a waveform corresponding to synthesized audio generated from concatenated phonetic units;
 - means for displaying parameters corresponding to at least one of the phonetic units;
 - means for receiving an editing input from the user; and
 - means for adjusting the parameters in accordance with the editing input.